



KONGERIKET NORGE

The Kingdom of Norway

Bekreftelse på patentsøknad nr

Certification of patent application no

1999 6331

► Det bekreftes herved at vedheftede dokument er nøyaktig utskrift/kopi av ovennevnte søknad, som opprinnelig inngitt 1998.12.23

► *It is hereby certified that the annexed document is a true copy of the above-mentioned application, as originally filed on 1998.12.23*

2003.07.17

Freddy Strømmen

Freddy Strømmen
Seksjonsleder

Line Reum

Line Reum



PATENTSTYRET®

Styret for det industrielle rettsvern

1c

PATENTSTYRET

20.DES99 996331

20 DES. 1999

EK/KBN

17.12.1999

E10508

UTSKILT FRA SØKNAD nr. 19986133 av 23/12-1998

Preben Lexow
Fløensbakken 41A
5009 Bergen

Oppfinner:

Søkeren

Metoder for sekvensanalyse

Foreliggende oppfinnelse er utskilt fra patentsøknad 19986133 som omfatter en metode for DNA-sekvensering som inneholder følgende trinn:

Første trinn tar utgangspunkt i en ren DNA-populasjon bestående av DNA-sekvensen som skal sekvenseres. DNA-molekylene kuttes/brekkes på en uspesifikk måte slik at det dannes en populasjon med DNA-molekyler bestående av biter (heretter kalt DNA-biter) av den opprinnelige sekvensen.

Annet trinn består i å erstatte baseparene i DNA-bitene med 4 ulike DNA-sekvenser (heretter kalt DNA-fragmenter) som representerer hver av de fire basene adenin, cytosin, guanin og tymin. Der hvor det har vært basepar A-T, settes det altså inn "fragment A", C-G byttes ut med "fragment C" osv. Dermed genereres nye DNA-molekyler hvor den opprinnelige baserekkefølgen på f.eks. ACGTT... erstattes med fragment A - fragment C - fragment G osv. Lengden på disse fire DNA-fragmentene kan i prinsippet variere i lengde fra 2bp til flere hundre kbp (eller mer om ønskelig), alt etter behov. Tilsvarende kan DNA-fragmentene inneholde reportergener og annen biologisk informasjon eller kun bestå av sekvenser uten kjent biologisk funksjon.

I tredje trinn avleses rekkefølgen av de fire typene DNA-fragmenter for hvert enkelt DNA-molekyl. Dermed finner man baserekkefølgen i de opprinnelige DNA-bitene indirekte.

I fjerde trinn benytter et dataprogram overlappen mellom DNA-bitene til å sette sammen informasjonen fra trinn 3 til sekvensen på DNA-sekvensene som ble brukt som utgangspunkt.

Foreliggende oppfinnelse omhandler alternative måter for utførelsen av trinn 2 og 3; hvor konverteringen og avlesningen utføres på det samme underlaget.

Alternative konverteringsmekanismer på fast matrix

I den opprinnelige patentsøknaden er konverteringsmekanismene holdt adskilt fra avlesningsmekanismene. Det er imidlertid også mulig å la konverteringen og avlesningen foregå på det samme underlaget. Nedenfor følger eksempler på slike prosedyrer. Et viktig poeng med disse prosedyrene er at man i tillegg til å få basesammensetningen av korte områder (10-100bp, eller mer) også får informasjon om deres innbyrdes lokalisasjoner på større DNA-molekyler (opptil flere kb). Dette er av betydning for rekonstruksjonen i trinn 4 og metoden kan f.eks. brukes til å komplementere sekvensinformasjon som er utledet vha de tidligere omtalte alternativene.

Alternativ 1

Utgangspunktet for denne prosedyren er en PCR-reaksjon hvor man amplifiserer opp DNA-sekvensen som skal sekvenseres. Den ene primeren er biotinmerket i den ene enden slik at man kan feste DNA-molekylene til et streptavidinunderlag. Streptavidinen er arrangert i en tynn linje slik at DNA-molekylene er festet ved siden av hverandre på en rekke. Deretter behandles DNA-molekylene slik som illustrert i fig. 1

Et viktig poeng med avlesningsmetoden er at adapternes relative posisjoner varierer med den relative posisjonen til sekvensen som gjenkjennes. Dermed blir det lettere å rekonstruere sekvensen i siste trinn. På figur 1 ser man f.eks at adapteren ACG befinner seg lengst til venstre mens adapteren GCT lengst til høyre, noe som gjenspeiler disse sekvensenes innbyrdes posisjoner på polynukleotidet. Hvis man regner med en fluorescenshalo på omlag 100nm og at 100nm utstrukket DNA tilsvarer omtrent 40-50 bp betyr det at man sannsynligvis vil klare å plassere hver adapter med en presisjon på 10-50bp.

Videre må det poengteres at sekvensinformasjonen som avleses per DNA-molekyl kan gjøres langt større enn 3 bp som i figur 1. Ved å anvende konverteringsmetodene som jeg har beskrevet i den opprinnelige patentsøknaden av 23/12-98 samt ettersendte revisjoner, før DNA-molekylene festes til streptavidinunderlaget kan man sannsynligvis få sekvensinformasjon som tilsvarer 50 bp eller mer per DNA-molekyl som konverteres. Det er også mulig å bruke konverteringsmekanismen illustrert i fig. 2.

Til slutt må det understrekkes at avlesningspotensialet kan økes ved å avlese flere ulike DNA-sekvenser på samme avlesningsplate (fig. 3). F.eks. kan man vha PCR amplifisere opp et større antall gener med genspesifikke primere. Deretter lager man unike overheng på de oppamplifiserte gensekvensene. Det kan f.eks. gjøres ved å bruke ekstra lange primere i siste syklus eller legge inn kutester for restriksjons endonukleaser som kutter sjeldent, mm. Dermed kan man hybridisere genene til en DNA-chips hvor hvert kvadrat består av oligonukleotider som er spesifikke for de ulike genene. DNA-molekylene som tilsvarer GenA vil dermed hybridiseres til kvadrat A, GenB til kvadrat B, osv. Denne metoden er særlig egnet til massescreening av enkeltpersoners genomer da man har muligheten til å plukke ut de spesifikke områdene av et genom som er av f.eks. medisinsk interesse.

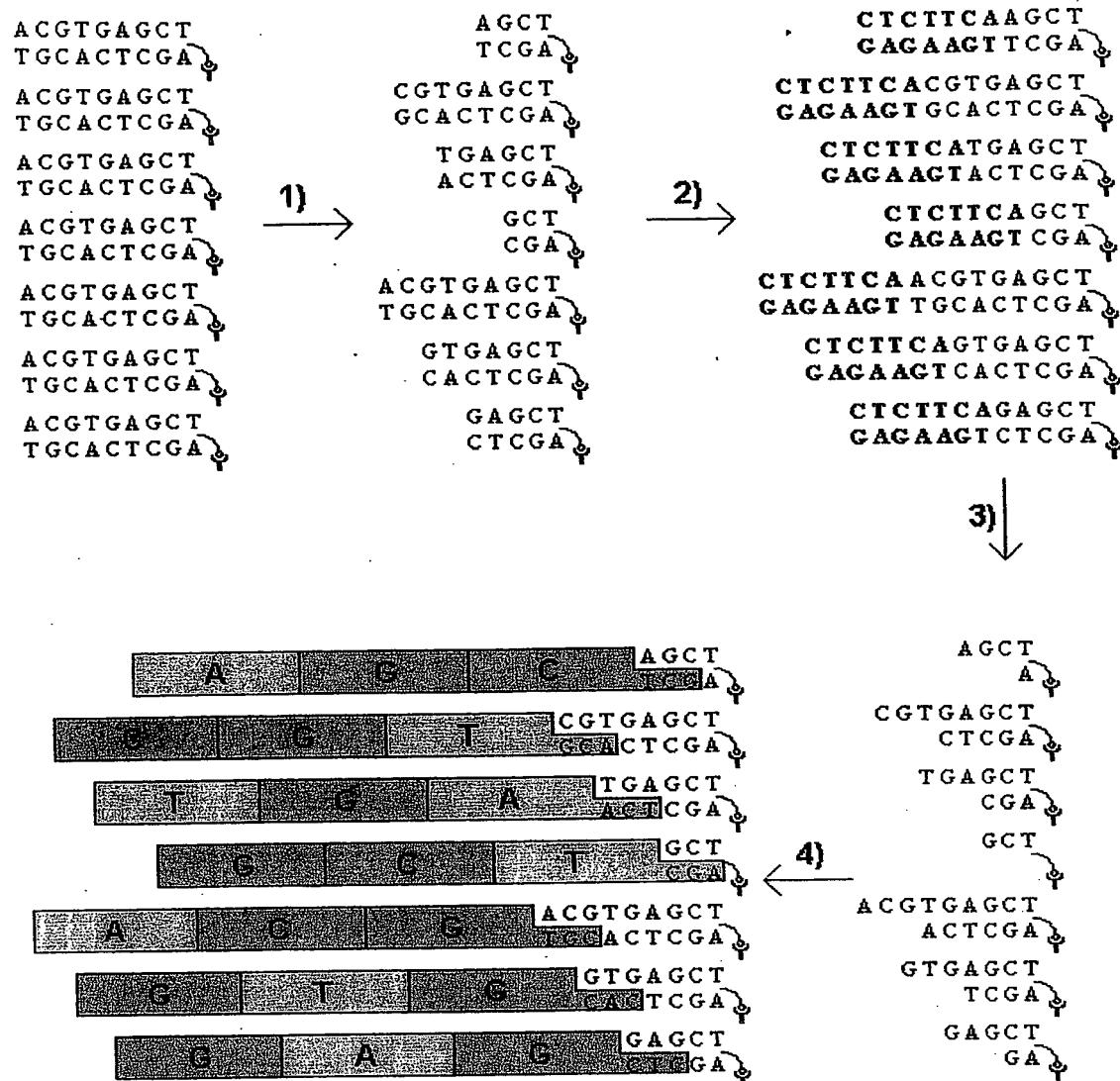


Fig.1 Prinsippet bak alternativ1. For å illustrere prinsippet er det her vist sekvenseringen av et ni basepar langt polynukleotid, ACGTGAGCT. I virkelighetenvil vil det være aktuelt å sekvensere mye lengre DNA-molekyler (opp til flere kb eller lengre). Etter at polynukleotidet er amplifisert opp og festet til et underlag på en rett linje behandles DNA-molekylene med DnaseI eller lignende slik at de kuttes tilfeldig (trinn 1)). I trinn 2) ligeres de kuttede endene med et polynukleotid som inneholder et bindingssett for en klasselII restriksjons endonuklelease som kutter utenfor sitt eget bindingssett (I dette tilfellet Earl). Når man tilsetter restriksjons endonukleasen i trinn 3) dannes det dermed overheng i polynukleotidene. I trinn 4) tilsettes adaptere som gjenkjerner og ligeres spesifikt til polynukleotid overhengene. Til det øverste polynukleotidet med overheng AGC ligeres det dermed en adapter med fragmentkombinasjonen AGC, osv. Deretter rettes DNA-molekylene ut ved hjelp av en væskestrøm, elektriskt felt eller lignende slik at de fluorescensmerkede adapterne kan avleses med en fluorescensscanner. Til slutt rekonstrueres sekvensen ved å aligne bitene med sekvensinformasjon. Legg forøvrig merke til at den relative posisjonen til hver adapter varierer etter hvor polynukleotidet ble kuttet av DnaseI. Dermed kan hver bit med sekvensinformasjon gies en relativ posisjon på polynukleotidet noe som gjør det lettere å rekonstruere sekvensen.

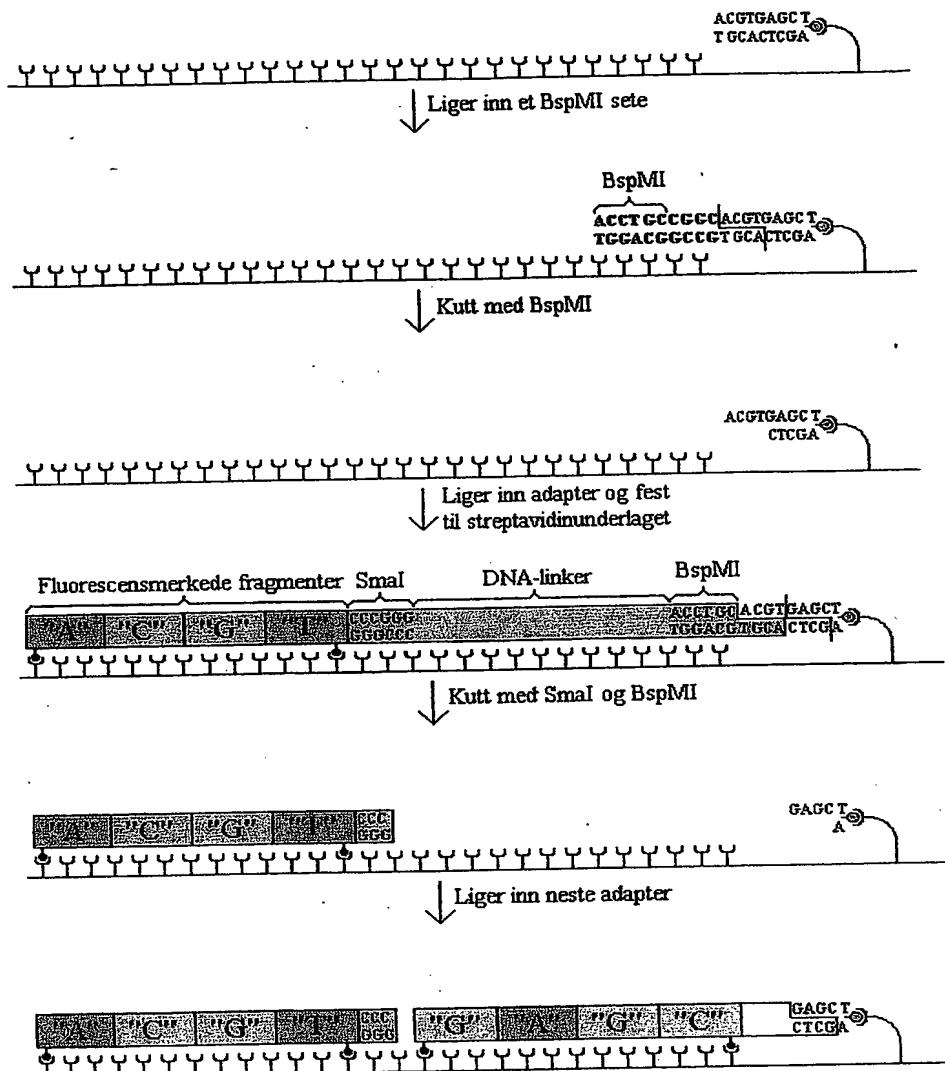
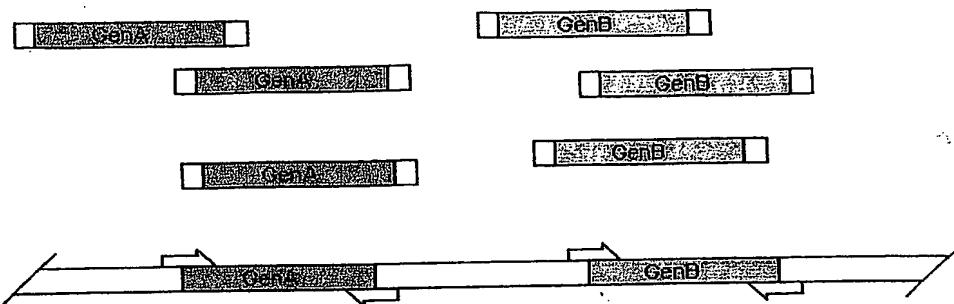


Fig. 2 . Prinsipp for å bygge opp lengre fragmentkjeder. DNA-molekylet som skal sekvenseres, ACGTGAGCT, festes med den ene enden til en streptavidindekt plate. Festmekanismen bør være en annen en streptavidin/biotin. Deretter ligeres DNA-molekylet med et polynukleotid som inneholder et bindingssete for en typeII restriksjons endonuklease med kuttssete utenfor bindingsssete (F.eks. BspMI som vist her). Når det i neste trinn kuttes med restriksjons endonukleasen dannes det et overheng med baser fra DNA-molekylet som skal sekvenseres. Deretter tilsettes det en løsning med ulike adaptere og ligase. På figuren ser man en adapter som har gjenkjent og bundet seg til overhenget ACGT. I tillegg til fluorescensmerkede fragmenter som motsvarer overhenget ACGT er det innkorporert to eller flere biotinmolekylet på adapteren. Når DNA-molekylet rettes ut ved hjelp av en væskestøm eller elektrisk felt er det dermed mulig å feste fragmentene til underlaget slik som vist på figuren. (Funksjonen til DNA-linkeren er å få fragmentene festet et stykke unna det DNA-molekylet som skal sekvenseres. Dermed blir det plass til en ny adapter i neste trinn). Deretter kutter man med SmaI og BspMI slik at DNA-linkeren fjernes samtidig som det dannes et nytt overheng bestående av de fire neste baseparene på DNA-molekylet. Dermed kan man innliggere en ny adapter med fluorescensmerkede fragmenter. Eneste forskjellen er at denne adapteren ikke inneholder en DNA-linker. Dermed blir de fluorescensmerkede adapterne festet til en ny posisjon på streptavidinunderlaget. Ved å bruke DNA-linkere med ulike lengder er det i prinsippet mulig å gjøre svært mange slike konverteringssykler etter hver andre.

1)



2)

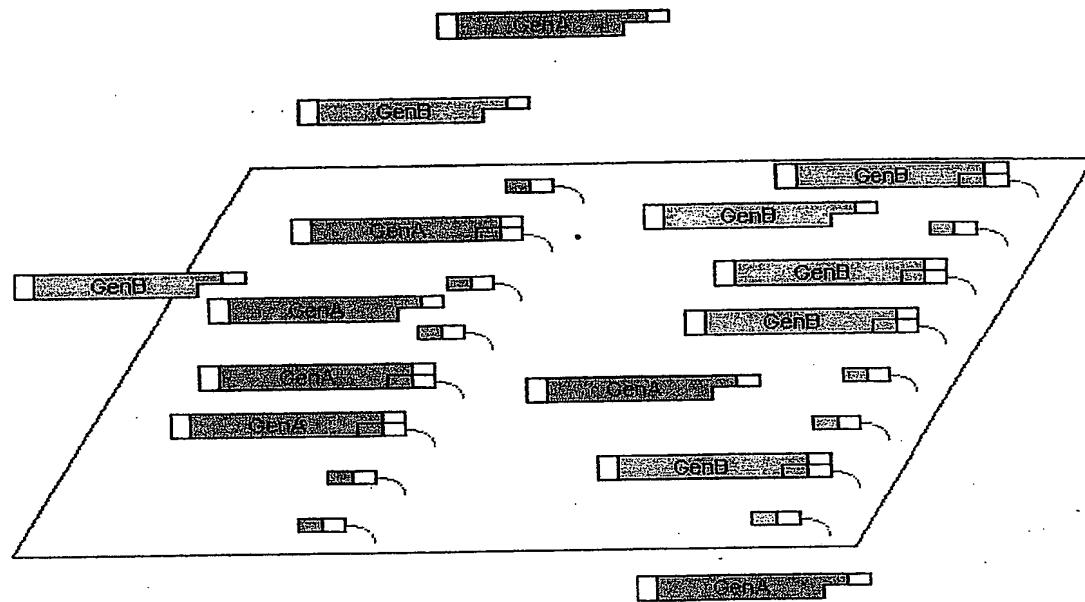


Fig. 3 Metode for parallell sekvensering av gener ved hjelp av alternativ1. 1) Genene som ønskes sekvensert amplifiseres opp fra genomisk DNA ved hjelp av genspesifikke primere. 2) Det genereres overheng som er spesifikke for hvert gen før genene hybridiseres til oligonukleotider på en DNA-chips. DNA-chipsen er laget slik at oligonukleotidene som er komplementære med genA overhengene befinner seg til venstre på avlesningsplaten, mens de genB komplementære befinner seg til høyre. I virkeligheten er det mulig å lage flere tusen ulike adresser på samme avlesningsplaten slik at det er mulig å sekvensere flere tusen gener parallelt.

Alternativ2

Utgangspunktet for prosedyren er at DNA-molekylene som skal sekvenseres kuttes opp i molekyler på noen få kb eller lengre. Deretter innkorporerer man biotin i DNA-molekylene slik at det gjennomsnittlig er baser med biotin med f.eks. noen hundre basepairs mellomrom (evt mer eller mindre avhengig av behov). DNA-molekylene festes deretter med den ene enden til en plate som er dekket med streptavidin. Festemekanismen for endene bør være en annen enn streptavidin/biotin. Molekylene rettes så ut ved hjelp av en væskestørm, elektrisk felt, eller annet. Til slutt forankres DNA-molekylene til underlaget ved å tilsette en reaksjonsløsning som forårsaker biotin-streptavidin binding og behandles videre som beskrevet i figur 4.

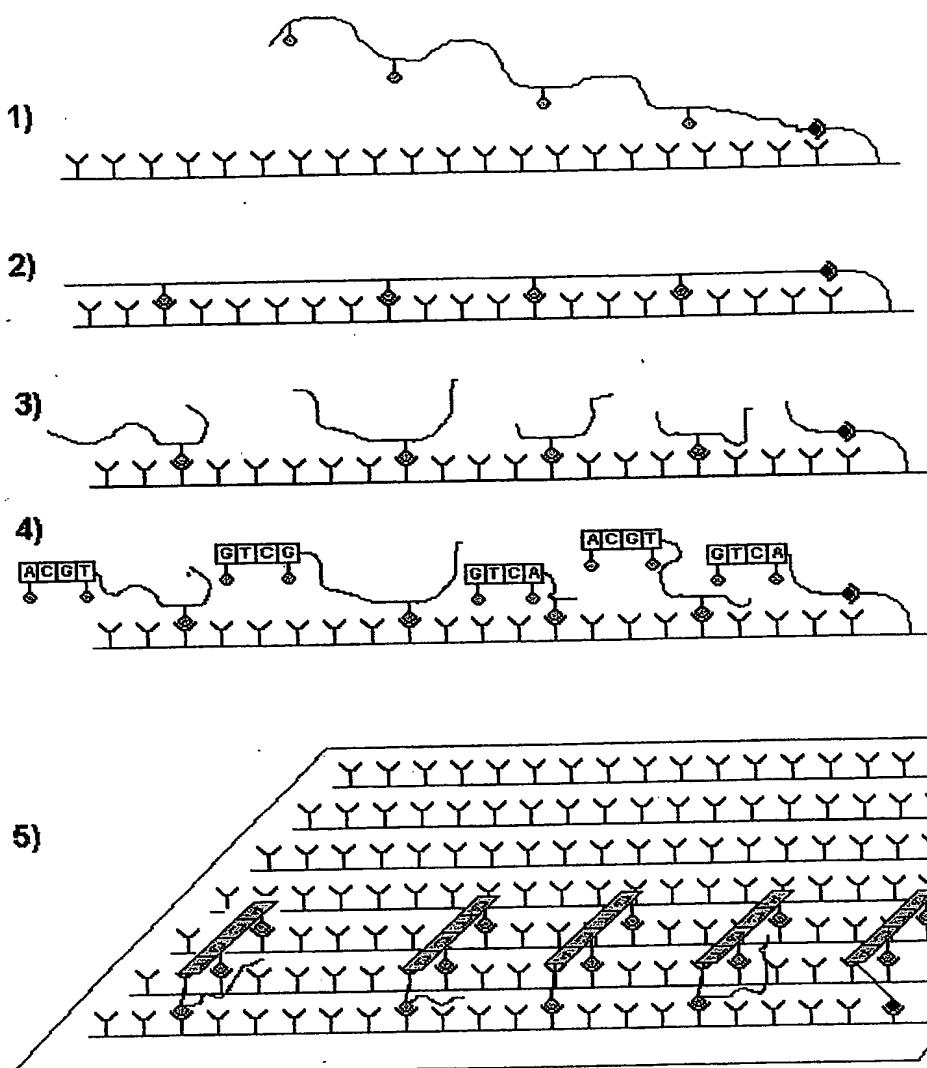
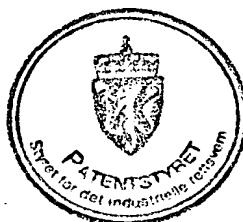


Fig.4 Behandlingen av DNA-molekylet initieres med at den ene enden festes til underlaget (1). DNA-molekylet rettes deretter ut ved hjelp av en væskestrøm, elektrisk felt eller lignende og forankres deretter til det streptavidindekte underlaget ved hjelp av de innkorporerte biotinmolekylene (2). Deretter kuttes DNA-molekylet med DnaseI eller lignende (3) før man ligerer de frie endene med preadAPTERe som inneholder bindingsseter for type II restriksjons endonukleaser (ikke vist). Ved deretter å kutte med den respektive endonukleasen dannes det overheng som ligeres med adaptere med sekvensinformasjon. Til ACGT-overhenget ligeres det dermed en adapter med fragmentkombinasjonen ACGT, osv (4). I siste trinn bruker man en væskestrøm, elektrisk felt eller lignende til å rette ut DNA-adapterne i en retning 90 grader på retninga til DNA-molekylet som skal sekvenseres, før de forankres ved hjelp av biotin/streptavidin systemet. Siden alle DNA-molekylene er forankret til underlaget kan prosessen gjentaes til man får det ønskede antallet basepar konvertert. Legg for øvrig merke til at de relative avstandene mellom adapterne tilsvarer de innbyrdes avstandene til sekvensbitene har i DNA-molekylet som skal sekvenseres. Videre må det nevnes at man selvfølgelig kan sekvensere svært mange DNA-molekyler parallelt på en avlesningsplate.



P a t e n t k r a v

1.

Fremgangsmåte for sekvensanalyse, karakterisert ved beskrivelsen.



Sammendrag

O. nr. E10508

Metode for konvertering og avlesning til bruk ved sekvensanalyse.

